

Süni intellekt dövründə dil: filoloji tədqiqatlarda yeni metodlar, imkanlar və etik çağırışlar

Samirət Məmmədova 

Xülasə. Məqalədə süni intellektin, xüsusilə böyük dil modellərinin filoloji tədqiqatların metodologiyasına təsiri araşdırılır. Tədqiqatın məqsədi dil texnologiyalarının mətn emalı, korpus təhlili, semantik şərh və çoxdilli araşdırmalarda yaratdığı imkanları sistemləşdirmək, eyni zamanda onların epistemoloji və etik məhdudiyyətlərini müəyyənləşdirməkdir. Məqsədli seçilmiş 12 elmi mənbə keyfiyyət məzmun təhlili vasitəsilə araşdırılmışdır. Təhlil nəticəsində dörd əsas istiqamət müəyyən edilmişdir: tədqiqat prosesinin sürətləndirilməsi, genişmiqyaslı dil məlumatlarının emalı, azresurslu dillərin rəqəmsal görünürsüzlüyünün artırılması və yeni etik-etibarlılıq problemlərinin yaranması. Nəticələr göstərir ki, süni intellekt filoloqu əvəz edən müstəqil şərhçi deyil, tədqiqatçının nəzarəti altında işlədilən köməkçi analitik vasitə kimi daha səmərəlidir. Məqalədə insan nəzarəti, mənbələrin yoxlanılması, dil müxtəlifliyinin qorunması və metodoloji şəffaflıq əsas tətbiq prinsipləri kimi təklif olunur.

Açar sözlər: süni intellekt, böyük dil modelləri, filologiya, təbii dilin emalı, korpus dilçiliyi, azresurslu dillər, akademik etika

Bakı Dövlət Universiteti, Bakı, Azərbaycan

E-poçt: filologiya_dekanliq@mail.ru

Daxil oldu: 11 Mart 2026; Qəbul edildi: 12 May 2026; Onlayn dərc edildi: 5 Sentyabr 2026

© Müəllif(lər) 2026. Bu açıq girişli məqalə Creative Commons Attribution–NonCommercial 4.0 Beynəlxalq Lisenziyasının (CC BY-NC 4.0) şərtlərinə uyğun olaraq yayımlanmışdır.

Language in the Age of Artificial Intelligence: New Methods, Opportunities, and Ethical Challenges in Philological Research

Samirat Mammadova 

Abstract. This article examines the influence of artificial intelligence, particularly large language models, on the methodology of philological research. It aims to systematize the opportunities created by language technologies in text processing, corpus analysis, semantic interpretation, and multilingual research while identifying their epistemological and ethical limitations. A purposive corpus of 12 scholarly sources was examined through qualitative content analysis. Four principal themes emerged: acceleration of research processes, processing of large-scale linguistic data, increased digital visibility of low-resource languages, and new problems concerning ethics and reliability. The findings indicate that artificial intelligence is most effective not as an autonomous interpreter replacing the philologist but as an auxiliary analytical instrument operating under researcher supervision. Human oversight, source verification, protection of linguistic diversity, and methodological transparency are proposed as the principal conditions for responsible use.

Keywords: *artificial intelligence, large language models, philology, natural language processing, corpus linguistics, low-resource languages, academic ethics*

Baku State University, Baku, Azerbaijan

E-mail: filologiya_dekanliq@mail.ru

Received: 11 March 2026; Accepted: 12 May 2026; Published online: 5 September 2026

© The Author(s) 2026. This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Giriş

Rəqəmsallaşmanın sürətlənməsi dil haqqında elmi biliklərin əldə olunması, saxlanması və təhlili üsullarını əsaslı şəkildə dəyişdirmişdir. Korpus bazaları, avtomatik morfoloji analizatorlar, maşın tərcüməsi sistemləri və böyük dil modelləri filoloji tədqiqatlarda əvvəllər uzun müddət tələb edən əməliyyatların qısa vaxt ərzində yerinə yetirilməsinə imkan verir. Xüsusilə transformer arxitekturasının tətbiqi söz və cümlələr arasındakı uzaq əlaqələrin modelləşdirilməsini gücləndirmiş, bununla da mətnlərin avtomatik təsnifatı, semantik yaxınlığın müəyyənəşdirilməsi və çoxdilli məlumatların işlənməsi üçün yeni texniki zəmin yaratmışdır (Vaswani et al., 2017). BERT kimi kontekstual dil modelləri sözün mənasını yalnız lüğəvi vahid kimi deyil, daxil olduğu mətn mühitinə əsasən modelləşdirməyə yönəlmişdir (Devlin et al., 2019).

Bu dəyişiklik filologiya üçün yalnız texniki yenilik deyil. Dil insanın tarixi yaddaşını, sosial münasibətlərini, mədəni kateqoriyalarını və dünyagörüşünü daşıdığına görə avtomatik dil təhlilinin nəticələri həm nəzəri, həm də etik baxımdan qiymətləndirilməlidir. Hesablama sistemləri böyük həcmli məlumatda təkrarlanan strukturları sürətlə aşkar etsə də, onların təqdim etdiyi nəticə dil faktının tarixi və mədəni izahını öz-özünə təmin etmir. Manning (2015) dərin öyrənmənin hesablama dilçiliyində güclü nəticələr yaratdığını göstərməklə yanaşı, dilçilik biliklərinin və nəzəri şərhin əhəmiyyətinin aradan qalxmadığını vurğulayır.

Böyük dil modellərinin geniş yayılması ilə məsələ daha mürəkkəb xarakter almışdır. Bu sistemlər akademik üsluba uyğun və zahirən inandırıcı mətn yarada bilir, lakin yanlış fakt, mövcud olmayan istinad və ya mənbədə ifadə edilməyən nəticə təqdim edə bilər. Bender et al. (2021) dil modelinin yaratdığı səlis formanın həqiqi məna və anlayışla eyniləşdirilməsinin metodoloji risklərini göstərirlər. Floridi və Chiriatti (2020) də GPT tipli modellərin imkanlarını onların faktiki, semantik və etik məhdudiyyətləri ilə birlikdə qiymətləndirməyin zəruriliyini qeyd edirlər.

Mövcud tədqiqatlarda süni intellekt əsasən kompüter elmləri, təhsil və kommunikasiya kontekstində araşdırılır. Filoloji tədqiqat baxımından isə üç məsələni eyni çərçivədə birləşdirən yanaşmaya ehtiyac vardır: yeni metodların elmi faydası, dil müxtəlifliyinə təsir və nəticələrin etibarlılığı. Bu məqalənin məqsədi süni intellekt əsaslı dil texnologiyalarının filologiyada tətbiq imkanlarını sistemləşdirmək, onların metodoloji və etik risklərini müəyyənəşdirmək və məsuliyyətli istifadə üçün praktik prinsiplər irəli sürməkdir.

Tədqiqat aşağıdakı suallara cavab axtarır: (1) süni intellekt filoloji tədqiqatın hansı mərhələlərini dəyişdirir; (2) bu texnologiyalar azresurslu və türk dilləri üçün hansı imkanları yaradır; (3) avtomatik nəticələrin etibarlılığını zəiflədən əsas risklər hansılardır; (4) filoloqun tədqiqat prosesində rolu necə qorunmalıdır?

Metodlar

Tədqiqat konseptual icmal və keyfiyyət məzmun təhlili əsasında aparılmışdır. Təhlil korpusu süni intellekt, təbii dilin emalı, böyük dil modelləri, dil müxtəlifliyi, azresurslu dillər və etik risklər üzrə məqsədli seçilmiş 12 elmi mənbədən təşkil edilmişdir. Mənbələr 2009–2023-cü illəri əhatə edir və resenziyalı jurnal məqalələri, beynəlxalq konfrans materialları, elmi kitab və texniki hesabatlardan ibarətdir. Məqsəd statistik meta-analiz aparmaq deyil, filoloji tədqiqat üçün təkrarlanan metodoloji mövqeləri və problem sahələrini müəyyənləşdirmək olmuşdur.

Mənbələrin seçilməsində dörd meyar tətbiq edilmişdir: (a) dilin avtomatik emalı və ya böyük dil modelləri ilə birbaşa əlaqə; (b) filoloji və dilçilik tədqiqatına tətbiq oluna bilən metodoloji məzmun; (c) dil müxtəlifliyi, qərəz və etibarlılıq məsələlərinə toxunma; (d) nəşr və müəllif məlumatlarının yoxlanıla bilməsi. Təhlil vahidi hər mənbədə irəli sürülən əsas metod, imkan, risk və tövsiyə olmuşdur.

Kodlaşdırma iki mərhələdə həyata keçirilmişdir. Birinci mərhələdə mənbələrdən tədqiqat alətləri, dil əhatəsi, insan nəzarəti, məlumat qərəzi, şəffaflıq və etik risklərə aid fikirlər çıxarılmışdır. İkinci mərhələdə oxşar kodlar birləşdirilərək dörd tematik kateqoriya yaradılmışdır: metodoloji sürət və miqyas; kontekstual-semantik təhlil; dil müxtəlifliyi və azresurslu dillər; etibarlılıq və etik məsuliyyət. Nəticələrin şərhində texniki məhsuldarlıqla filoloji izahın eyni anlayış olmadığı prinsipi əsas götürülmüşdür.

Nəticələr

Mənbələrin tematik təhlili süni intellektin filoloji tədqiqata təsirini dörd əsas istiqamətdə göstərmişdir. Bu istiqamətlər bir-birindən tam ayrılmır; texniki imkanların artması eyni zamanda yeni yoxlama və məsuliyyət mexanizmlərinə ehtiyac yaradır.

Cədvəl 1

Süni intellektin filoloji tədqiqatlarda əsas imkan və məhdudiyyətləri

Tematik istiqamət	Əsas imkan	Əsas məhdudiyyət
Sürət və miqyas	Böyük mətn korpuslarının axtarışı, təsnifatı və ilkin kodlaşdırılması	Sürətli nəticənin avtomatik olaraq düzgün nəticə sayılması
Semantik təhlil	Kontekstual uyğunluqların, mövzu və mənə əlaqələrinin müəyyənləşdirilməsi	Tarixi, ironik və mədəni mənaların səthi şərh
Dil müxtəlifliyi	Azresurslu dillər üçün korpus, tərcümə və rəqəmsal resursların yaradılması	Məlumat azlığı və dominant dillərin struktur qərəzi
Etika və etibarlılıq	Tədqiqat prosesinin dəstəklənməsi və variantların müqayisəsi	Uydurma istinad, qərəz, müəlliflik və şəffaflıq problemləri

Mənbə. Müəllif tərəfindən hazırlanmışdır.

Metodoloji sürət və tədqiqat miqyasının genişlənməsi

Birinci nəticə süni intellektin tədqiqatın hazırlıq və ilkin analiz mərhələlərini sürətləndirməsidir. Rəqəmsal mətnlərin tokenləşdirilməsi, lemmatizasiyası, nitq hissələrinə görə işarələnməsi, tezlik siyahılarının hazırlanması və açar sözlərin çıxarılması böyük korpusların tədqiqatını mümkün edir. Bird et al. (2009) təbii dilin emalı alətlərinin mətn üzərində formal əməliyyatları təkrarlana bilən prosedurlara çevirdiyini göstərir. Beləliklə, filoloq ayrı-ayrı nümunələri əl ilə toplamaq əvəzinə, daha geniş materialda qanunauyğunluqları müəyyən edə bilər. Süni intellekt və təbii dilin emalı üsulları böyük mətn korpuslarında yeni leksik vahidlərin müəyyənləşdirilməsinə, onların yaranma modellərinin araşdırılmasına və semantik inkişafının izlənməsinə də imkan verir (Seidova et al., 2026).

Bununla belə, avtomatik seçim tədqiqatçının qərarını aradan qaldırmır. Korpusun tərkibi, mətnlərin dövrü, janrı, müəllifliyi və orfoqrafik norması nəzərə alınmadıqda yüksək tezlik yanlış ümumiləşdirməyə səbəb ola bilər. Süni intellekt nəticəni sürətləndirir, lakin tədqiqat obyektinin sərhədlərini filoloqun əvəzinə müəyyənləşdirmir.

Kontekstual və semantik təhlil imkanları

İkinci nəticə kontekstual dil modellərinin semantik oxşarlıq və mətn əlaqələrinin aşkarlanmasında yeni imkanlar yaratmasıdır. Transformer əsaslı modellər sözün müxtəlif kontekstlərdə qazandığı mənaları müqayisə etməyə, mətnləri mövzu üzrə qruplaşdırmağa və diskursdakı leksik əlaqələri müəyyənləşdirməyə kömək edir (Devlin et al., 2019; Vaswani et al., 2017). Bu imkan terminoloji dəyişmələrin, sinonim sıralarının, metaforik istifadənin və üslubi fərqlərin geniş korpus üzrə ilkin xəritələşdirilməsində faydalıdır. Korpus və hesablama dilçiliyi metodlarının frazeoloji vahidlərin tezliyini, kollokasiya xüsusiyyətlərini və zamanla dəyişən mənalarını müəyyənləşdirmək üçün də faydalı olduğu göstərilir (Gasimov, 2026).

Lakin statistik yaxınlıq filoloji mənənin tam ekvivalenti deyildir. Tarixi mətnlərdə arxaizm, müəllif ironiyası, intertekstuallıq və mədəni işarələr modelin təlim məlumatında kifayət qədər təmsil olunmaya bilər. Buna görə avtomatik nəticə yekun şərh kimi deyil, tədqiqatçının mətnə qayıdaraq yoxlamalı olduğu hipotez kimi qəbul edilməlidir. Modelin verdiyi kateqoriya nümunə cümlələr və mətnin bütöv konteksti ilə təsdiqlənməlidir.

Dil müxtəlifliyi və azresurslu dillər

Üçüncü nəticə süni intellektin azresurslu dillər üçün həm imkan, həm də bərabərsizlik yaratmasıdır. Joshi et al. (2020) təbii dilin emalı sahəsində dillərin qeyri-bərabər təmsil olunduğunu, tədqiqat və texnoloji resursların əsasən az sayda dominant dil ətrafında cəmləşdiyini göstərir. Bender (2019) dilçilik və dil texnologiyası araşdırmalarında öyrənilən dilin konkret adlandırılmasını metodoloji məsuliyyət kimi irəli sürür. “Dil” adı altında yalnız ingilis dilinə əsaslanan nəticələrin universallaşdırılması digər dillərin struktur xüsusiyyətlərini görünməz edə bilər.

Türk dillərinə dair genişmiqyaslı maşın tərcüməsi tədqiqatı paralel məlumatın və birgə modelləşdirmənin qohum dillər üçün praktik faydasını nümayiş etdirir (Mirzaxalov et al., 2021). Azərbaycan dili baxımından bu yanaşma milli korpusların genişləndirilməsi, tarixi yazı nümunələrinin rəqəmsallaşdırılması, dialekt materiallarının işarələnməsi və açıq linqvistik resursların yaradılması üçün əhəmiyyətlidir. Bununla yanaşı, türk dillərinin hər birinin fonetik, morfoloji və semantik xüsusiyyətləri ayrıca nəzərə alınmalıdır; qohumluq avtomatik eynilik kimi qəbul edilə bilməz.

Etibarlılıq, qərəz və etik məsuliyyət

Dördüncü nəticə böyük dil modellərinin elmi etibarlılıq və etik məsuliyyətlə bağlı yeni problemlər yaratmasıdır. Dil modeli ehtimal əsasında növbəti dil vahidini seçdiyinə görə səlis, lakin faktiki cəhətdən yanlış mətn yarada bilər. OpenAI (2023) GPT-4-ün imkanları ilə yanaşı, hallüsinasiya, qərəz və etibarlılıq məhdudiyyətlərini də qeyd edir. Filoloji araşdırmada bu risk xüsusilə sitatların, etimoloji məlumatların, əsər adlarının və bibliografik göstəricilərin dəqiqliyinə təsir edir.

Blodgett et al. (2020) dil texnologiyalarında qərəzin yalnız texniki xəta deyil, sosial və normativ məsələ olduğunu göstərir. Weidinger et al. (2022) dil modellərinin risklərini ayrı-seçkilik, yanlış məlumat, zərərli məzmun və insan–kompüter qarşılıqlı əlaqəsi daxil olmaqla geniş taksonomiya üzrə qruplaşdırırlar. Buna əsasən filoloji tədqiqatda modelin nəticəsi müəllif, nəşr, səhifə və DOI kimi ilkin mənbə göstəriciləri üzrə ayrıca yoxlanılmalıdır. Süni intellektdən istifadə faktı metod bölünməsinə və ya jurnalın tələb etdiyi açıqlamada şəffaf şəkildə bildirilməlidir.

Müzakirə

Nəticələr süni intellektin filologiyada ikiqat funksiyaya malik olduğunu göstərir. Rəqəmsal kommunikasiya mühitində yeni leksik formaların, qısaltmaların və akronimlərin sürətlə yayılması texnoloji inkişafın dilin leksik-semantik sisteminə təsirini ayrıca araşdırmağı zəruri edir (Guliyeva, 2026). Bir tərəfdən, o, böyük həcmli mətn məlumatını emal edir, görünməyən nümunələri üzə çıxarır və çoxdilli müqayisəni asanlaşdırır. Digər tərəfdən, modelin nəticələrinin mənbədən və tarixi-mədəni kontekstdən uzaqlaşması elmi şərhin keyfiyyətini zəiflədə bilər. Bu iki funksiya arasında tarazlığın yaradılması üçün “insan nəzarətli filoloji təhlil” yanaşması məqsədəuyğundur.

Bu yanaşmada tədqiqatçı problemi və korpusu müəyyənləşdirir, süni intellekt ilkin emal və alternativ təsnifatlar təqdim edir, filoloq isə nəticələri orijinal mətn və etibarlı mənbələrlə yoxlayır. Son qərar statistik uyğunluğa deyil, dil faktının struktur, tarixi, pragmatik və mədəni izahına əsaslanır. Belə iş bölgüsü süni intellekti müəllif və ya avtonom ekspert mövqeyindən çıxararaq metodoloji alət statusunda saxlayır.

Təhlil həmçinin göstərir ki, azresurslu dillər üçün məsuliyyətli süni intellekt siyasəti yalnız hazır modellərdən istifadə ilə məhdudlaşmamalıdır. Keyfiyyətli milli korpusların, açıq lüğət bazalarının, morfoloji işarələnmiş mətnlərin və dialekt arxivlərinin yaradılması uzunmüddətli prioritet olmalıdır. Əks halda Azərbaycan dili üzrə avtomatik sistemlər dominant dillərdən ötürülmüş kateqoriyalara və natamam məlumatlara əsaslanacaqdır. Dil resurslarının yaradılmasında müəllif hüquqları, şəxsi məlumatların qorunması və mətn sahibinin razılığı nəzərə alınmalıdır. Məqalənin məhdudiyyəti mənbə korpusunun məqsədli və nisbətən kiçik seçilməsidir. Buna görə nəticələr bütün filoloji alt-sahələr üzrə statistik ümumiləşdirmə kimi deyil, konseptual çərçivə kimi qəbul edilməlidir. Gələcək tədqiqatlarda Azərbaycan dilində hazırlanmış korpuslar üzərində müxtəlif dil modellərinin leksik-semantik, morfoloji və pragmatik tapşırıqlarda dəqiqliyi empirik şəkildə müqayisə edilə bilər.

Nəticə

Süni intellekt filoloji tədqiqatların metodoloji imkanlarını genişləndirərək mətnlərin sürətli emalına, böyük korpusların təhlilinə, semantik əlaqələrin ilkin müəyyənləşdirilməsinə və azresurslu dillər üçün yeni rəqəmsal resursların yaradılmasına şərait yaradır. Bununla belə, səlis yaradılmış mətn elmi doğruluğun təminatı deyildir. Uydurma istinadlar, məlumat qərəzi, dominant dillərin üstünlüyü və kontekstin itirilməsi filoloji araşdırma üçün əsas risklərdir. Məqalənin əsas qənaəti budur ki, süni intellekt filoloqu əvəz edən subyekt kimi deyil, insan nəzarəti altında fəaliyyət göstərən köməkçi analitik vasitə kimi tətbiq edilməlidir. Məsuliyyətli istifadə üçün dörd prinsip vacibdir: mənbələrin ilkin nəşrlər üzrə yoxlanılması; istifadə edilən model və prosedurun şəffaf göstərilməsi; dil və mədəni kontekstin filoloq tərəfindən şərh olunması; azresurslu dillərin məlumat bazalarının məqsədli şəkildə inkişaf etdirilməsi. Bu prinsiplər texnoloji yeniliyi elmi dürüstlük və dil müxtəlifliyinin qorunması ilə uzlaşdırmağa imkan verir.

Ədəbiyyat

- Bender, E. M. (2019). The #BenderRule: On naming the languages we study and why it matters. *Communications of the ACM*, 62(10), 33–35. <https://doi.org/10.1145/3338097>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the Natural Language Toolkit*. O'Reilly Media.

- Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of “bias” in NLP. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5454–5476). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.485>
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2021). *On the opportunities and risks of foundation models*. arXiv. <https://doi.org/10.48550/arXiv.2108.07258>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (Vol. 1, pp. 4171–4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Gasimov, A. (2026). Semantic functions of phraseological units in modern English. *Scientific Research*, 6(2), 83–89. <https://doi.org/10.36719/2789-6919/55/83-89>
- Guliyeva, A. (2026). Acronyms and abbreviations in modern English language. *Journal of Philological Studies*, 1(1), 34–40.
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30, 681–694. <https://doi.org/10.1007/s11023-020-09548-1>
- Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 6282–6293). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.560>
- Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Pearson.
- Seidova, E., Babayev, J., Sadikhova, S., Aghayeva, G., & Mammadova, Z. (2026). AI-powered lexical innovation: Modeling neologisms with NLP. *Scientific Work*, 20(3), 171–179. <https://doi.org/10.36719/2663-4619/127/171-179>
- Manning, C. D. (2015). Computational linguistics and deep learning. *Computational Linguistics*, 41(4), 701–707. https://doi.org/10.1162/COLI_a_00239
- Mirzakhlov, J., Babu, A., Ataman, D., Kariev, S., Tyers, F. M., Abduraufov, O., Hajili, M., Ivanova, S., Khaytbaev, A., Laverghetta, A., Moydinboyev, B., Onal, E., Pulatova, S., Wahab, A., Firat, O., & Chellappan, S. (2021). A large-scale study of machine translation in Turkic languages. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 5876–5890). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.emnlp-main.474>
- Nivre, J., de Marneffe, M.-C., Ginter, F., Hajič, J., Manning, C. D., Pyysalo, S., Schuster, S., Tyers, F., & Zeman, D. (2020). Universal Dependencies v2: An evergrowing multilingual treebank collection. In *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 4034–4043). European Language Resources Association.
- OpenAI. (2023). GPT-4 technical report. arXiv. <https://doi.org/10.48550/arXiv.2303.08774>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998–6008). Curran Associates.
- Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Bihane, A., Hendricks, L. A., Rimell, L., Isaac, W., ... Gabriel, I. (2022). Taxonomy of risks posed by language models. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 214–229). Association for Computing Machinery. <https://doi.org/10.1145/3531146.3533088>